

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Office

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification⁶: C07K 16/00, A61K 39/395, G01N 33/577	A1	(11) International Publication Number: WO 96/02574 (43) International Publication Date: 1 February 1996 (01.02.96)
(21) International Application No.: PCT/EP95/02626 (22) International Filing Date: 6 July 1995 (06.07.95) (30) Priority Data: P 44 25 115.7 15 July 1994 (15.07.94) DE (71) Applicant (for all contracting states with the exception of the US): BOEHRINGER MANNHEIM GMBH [DE/DE]; Sandhofer Strasse 116, D-68305 Mannheim (DE). (72) Inventor; and (75) Inventor/Applicant (for US only): STEIPE, Boris [DE/DE]; Unterbrunnerstrasse 10, D-82131 Gauting (DE). STEINBACHER, Stefan [DE/DE]; Tölzerstrasse 8, D-82661 Lengries (DE).		(81) Contracting States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT, UA, UG, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG). Published: <i>With an international search report. Before the expiration of the period allowed for claim changes. The patent will be republished in the event of changes.</i>
(54) Title: PROCESS FOR MODIFYING THE STABILITY OF ANTIBODIES (57) Abstract The invention concerns a process for preparing a functional derivative or fragment of immunoglobulin or a functional immunoglobulin having improved stability in a eukaryotic or prokaryotic organism after transformation with an expression vector containing a recombinant gene which codes for said immunoglobulin derivative or fragment. The invention is characterized in that a modified gene is used in which, in contrast to the non-modified gene, at least one codon which codes for a more common amino acid is replaced. In this way active antibodies which are disulphide bridge-free can also be prepared. In a variant of the process, antibodies can also be adequately destabilized.		

Process for Modifying the Stability of Antibodies

The invention relates to a process for modifying the stability of antibodies (AB) and the use thereof, particularly in diagnostic procedures and therapeutics.

AB biotechnology is a rapidly expanding field with a focus on diagnostic procedures (*in vitro*: e.g., antigen identification; *in vivo*: e.g., imaging), treatment (particularly humanized AB with a prolonged serum half-life and reduced immunogenicity), and toxicology (e.g., anti-digoxin antibodies as a specific antidote in cardiac glycoside overdose). Other fields of application are developing in the induction of transplant tolerance (e.g., by anti-CD4 AB), in immunotherapy (e.g., CAMPATH in non-Hodgkin's lymphoma), and in catalytic AB, which enable stereoselective and regioselective catalysis in particular.

Natural antibody sequences are not optimized for stability; genetically engineered sequence hybrids (e.g., humanized AB or single-chain Fv fragments) are frequently highly destabilized. The consequences of this can be, for example:

- difficult refolding
- denaturation: (I) decomposition and (II) immunogenicity even at 37°C *in vivo*
- reduced avidity
- aggregation and loss of activity during storage

The addition of proteins from the DNAJ protein family (EP-A 0 556 726) or the HSP90 protein family (EP-A 0 551 916) to stabilize antibodies in solutions is known, for example. By comparison, no process is known thus far whereby antibodies can be stabilized by selective mutation of the amino acid sequence. It is in fact theoretically possible to create numerous point mutations in antibodies and to screen these mutants for stability. It has become evident with other proteins, however, that only one of 10^3 to 10^4 mutants exhibits improved stability. Such screening processes are thereby very expensive and furthermore limited to proteins having

identifiable functions, such as enzymatic activity (Rollence, 1988; Chen, 1989; Turner, 1992; Risse, 1992; Arase, 1993).

The genes in the variable domains of immunoglobulins have changed variously over the course of their evolution through multiple gene duplications and mutations. They are optimized for the capacity for selective and high-affinity binding of antibodies (Tonegawa, 1983; Berek, 1988; French, 1989). During this process, the sequences coding for the domains mutated randomly and B cells exhibiting improved antigen binding were selected and propagated (Berek, 1993). Although optimization of antigen binding capacity plays a dominant role, the quality of an antibody depends on the sum of a multitude of factors, such as antigen affinity, domain stability, interaction between the heavy and light chains and between variable and constant domains, protease sensitivity, and the possibility of export and secretion of the antibodies out of the cell. Natural antibodies, accordingly, are not necessarily optimized for stability.

It is known from Frisch (1994) that a human V_k protein is destabilized after an exchange of cysteine 23, which prevents the formation of the cysteine 23/cysteine 88 disulfide bond. This destabilization can be partially reversed by exchanging tryptophan 32 and histidine. This is an accidental result, however, which moreover is no longer compatible with the science of discovery.

The reason for this is that the V_k -protein REI described by Frisch is not a V_k -domain fragment of a naturally occurring antibody but a protein which is overexpressed as such in a myeloma cell line. REI is a protein which differs greatly in its composition from V_k domains, which are fragments of naturally occurring antibodies. REI has, for example, atypical amino acids in positions 50 (E) and 92 (Q). Because of the spatial configuration of the amino acids, a salt bond can presumably form between E 50 and H 32 and a hydrogen bond between Q 92 and H 32. This type of hydrogen bond, which does not occur in natural antibodies, then stabilizes this V_k protein.

The object of the invention is to make available a process by which the stability of antibodies can be modified such that these antibodies can be selectively stabilized, destabilized, or restabilized after destabilizing measures, such as the removal of disulfide bonds, for example.

The subject of the invention is a process for preparing a stability-improved functional antibody, functional derivative, or fragment thereof in a eukaryotic or prokaryotic organism through transformation with an expression vector containing a recombinant gene, which codes for the indicated immunoglobulin, derivative, or fragment, characterized in that

- a) the gene of at least one of the variable domains of immunoglobulin is compared with the Consensus Tables 1–6 and the table is thereby selected which has the highest homology for this domain,
- b) at least one codon of an amino acid is replaced in the gene of this variable domain, namely
 - aa) if this amino acid is not listed in its position in said selected table, by a codon for one of the listed amino acids, and/or
 - bb) if this amino acid is listed in its position in the selected table, by a codon for one of the listed amino acids with a higher frequency,
- c) the prokaryotic or eukaryotic organism with the thus modified gene is transformed and the antibody, the fragment, or derivative with the desired activity is expressed.

If necessary, the antibody can be isolated from the organism by methods familiar to one skilled in the art and optionally purified.

In a preferred embodiment of the invention, the process is carried out such that

- a) in the gene of the variable domain of the human heavy chain at least one codon for an amino acid is replaced, namely,
 - aa) if this amino acid is not listed in its position in Table 1, by a codon for one of the listed amino acids, and/or

- ab) if this amino acid is listed in its position in Table 1, by a codon for one of the listed amino acids with a higher frequency,
- b) in the gene of the variable domain of the mouse heavy chain
 - ba) if this amino acid is not listed in its position in Table 2, by a codon for one of the listed amino acids, and/or
 - bb) if this amino acid is listed in its position in Table 2, by a codon for one of the listed amino acids with a higher frequency,
- c) in the gene of the variable domain of the human kappa light chain
 - ca) if this amino acid is not listed in its position in Table 3, by a codon for one of the listed amino acids, and/or
 - cb) if this amino acid is listed in its position in Table 3, by a codon for one of the listed amino acids with a higher frequency,
- d) in the gene of the variable domain of the mouse kappa light chain
 - da) if this amino acid is not listed in its position in Table 4, by a codon for one of the listed amino acids, and/or
 - db) if this amino acid is listed in its position in Table 4, by a codon for one of the listed amino acids with a higher frequency,
- e) in the gene of the variable domain of the human λ light chain
 - ea) if this amino acid is not listed in its position in Table 5, by a codon for one of the listed amino acids, and/or
 - eb) if this amino acid is listed in its position in Table 5, by a codon for one of the listed amino acids with a higher frequency,
- f) in the gene of the variable domain of the mouse λ light chain

- fa) if this amino acid is not listed in its position in Table 6, by a codon for one of the listed amino acids, and/or
- fb) if this amino acid is listed in its position in Table 6, by a codon for one of the listed amino acids with a higher frequency,
- g) and the prokaryotic or eukaryotic organism is transformed and the antibody, fragment, or derivative with the desired activity is expressed.

The novel process is used such that the antibody to be stabilized is first sequenced and the sequence of its domains is compared with the consensus sequences listed in Tables 1-6 or the sequences of Kabat (1991). The amino acid positions are defined by the maximum homology of sequences. Then, one or several codons can be modified according to the invention, expediently by mutagenesis. It becomes evident that even the selective exchange of a codon can cause a marked change in the stability of an antibody. Preferably, however, two, three, or more codons are modified. A top limit for the number of exchanges is reached when other properties of the antibody, important for the desired application purpose (e.g., affinity, protease stability, selectivity), are negatively affected.

The process will be elucidated with an example:

The amino acid positions are first determined by a sequence comparison (maximum homology) using Tables 1-6 or the tables of Kabat (1991).

For a human antibody the stability of which is not optimal, it is found that amino acid H is located at position 15 of the heavy chain. It can be derived from Table 1 that G or S is preferred for position 15. It is accordingly advantageous to replace H by S or especially preferably by G. If it is discovered that amino acid A is in position 16 in this antibody, then it is preferred to replace A by Q, R, or G. Apparently, it is especially advantageous to replace A by G.

If the antibody, e.g., after position 35, has an intron of one or two amino acids, then it is preferred to delete at least one of these amino acids (to replace 35 a/35 b with "-"). This also applies to other optional introns. The tables are thus to be interpreted such that the amino acids in positions designated by a, b, etc. (e.g., 35a, 35b) are deleted preferably for stabilizing the antibody (thus replaced by amino acids "-"). For position 100 b in Table 1, this means that for stabilization, for example, an amino acid not listed can be replaced by G or S. Preferably, however, this amino acid is deleted. It is likewise advantageous, however, to delete G or S at this position as well.

To stabilize an antibody by the novel process and still to retain its other properties, such as antigen affinity in particular, preferably amino acids are exchanged that do not interfere with these properties if possible. For this reason, it is preferred that no exchanges be made in the antigen-binding loops or CDRs.

The preparation of antibody derivatives and fragments can proceed by methods familiar to one skilled in the art for the preparation of recombinant proteins. Such processes are described, for example, in EP-B 0 125 023 and EP-B 0 120 694 and in S.L. Morrison et al. (1984).

For the preparation of antibodies modified as taught by the invention, the complete DNA of the variable domains can be synthesized, for example (via oligonucleotide synthesis, as described in Sinha et al., NAR 12 (1984), 4539-4557, for example). The oligonucleotides can be coupled by PCR as described, for example, in Innis, ed., *PCR Protocols*, Academic Press (1990) and Better et al., *J. Biol. Chem.* 267 (1992), 16712-16118. Cloning and expression proceeds via standard procedures as described, for example, in Ausubel et al., eds., *Current Protocols in Molecular Biology*, John Wiley and Sons, New York (1989), and in Robinson et al., *Hum. Antibod. Hybridomas* 2 (1991), 84-93. The specific antigen binding activity can be tested by a competitive test, for example, as described in Harlow et al., eds. *Antibodies: A Laboratory Manual*, Chapter 14, Cold Spring Harbor Laboratory, Cold Spring Harbor (1988), and Munson et al., *Anal. Biochem.* 407 (1980), 220-239.

Suitable host organisms are, for example, CHO cells, lymphocyte cell lines that produce no immunoglobulin, yeast, insect cells, and prokaryotes such as *E. coli*.

Another subject of the invention is a process in which the protein is isolated in a prokaryotic organism (e.g., *E. coli*) as denatured inclusion bodies and activated by processes familiar to one skilled in the art (cf., e.g., EP-A 0 364 926). In this regard, the activation can surprisingly also occur under reducing conditions.

Another subject of the invention is a process by which the antibody is stabilized as taught by the invention such that it arises biologically active with the desired activity in the cytosol and can be isolated therefrom directly and in active form.

The novel process improves the stability of antibodies and antibody fragments for all fields of application claimed above. Moreover, according to the invention new stable antibody variants can be prepared, which thus far could not be obtained in stable form, such as disulfide-bond-free AB or catalytic antibodies, which are suitable for use under nonphysiological conditions.

Catalytic antibodies and the use thereof are described, for example, in *CIBA Foundation Symposium on Catalytic Antibodies*, London, 1990, Vol. 159, Chadwick, D. J. and Marsh, J., eds., Wiley and Sons, Chichester.

Stabilized disulfide-bond-free antibodies are obtained by the replacement of the cysteines forming disulfide bonds by other amino acids and replacement of at least one, preferably two or more amino acids by stability-imparting amino acids.

Preferably, such antibodies are chimeric, humanized, nonhuman, or human antibodies assignable to β -lymphocyte expression (no REI proteins).

Another subject of the invention is a process for the preparation of nondisruptive destabilized antibodies, which can be used advantageously, for example, if rapid pharmacokinetics are

required. To obtain such antibodies, at least one amino acid exchange must consequently be carried out in a way opposite to that described above. This means that an amino acid with a higher frequency is replaced by an amino acid with a lower frequency.

Suitable as antibody fragments are, for example, Fab, Fab', F(ab')₂, single-chain antibodies, Fv, or individual variable domains. These fragments can also be coupled to other substances, for example, to immunotoxins.

The novel process is especially advantageous for improving the stability of single-chain F_v regions of antibodies, particularly, for improving the stability of single-chain immunotoxins. The light and heavy chains are bound to each other in various ways in such single-chain antigen-binding proteins. This binding is possible, for example, via a disulfide bond, via covalent bonds, or via a zinc complex bond. Such single-chain proteins and their coupling are described, for example, in Brinkmann et al., *P.N.A.S.* 89 (1992), 3075–3079 (Coupling via a peptide linker), and in Brinkmann et al., *P.N.A.S.* 90 (1993), 7536–7542 (Additional disulfide bonds). Other immunotoxins and bonding options are described in WO 91/09871, WO 91/12820, and WO 91/16069.

Another advantage of the invention is that scF_v (single-chain F_v, hybrid proteins of the V_H and V_L domains linked by an unstructured oligopeptide) are stable according to the invention and can be prepared in a less immunogenic form. The typically employed linker peptides (S.H. Jung, 1994; R. Glockshuber, 1990) of scFvs frequently lead to aggregation problems and are potential immunogens. The covalent linking of V_H and V_L domains, by comparison, can also occur via an intermolecular cystine bond; however, such additional cysteines thus far resulted in a considerable lowering of the folding yield because of the possibility of false disulfide bond formation. The conserved cysteines at positions 23/88 (light chain) and 22/99 (heavy chain) can be replaced by mutagenesis in the novel process, and the stability of the antibodies can be restored or improved by the novel process. Thereby, the false formation of disulfide bonds is

ruled out. The novel process is very important for this reason for the therapeutic utilization of recombinant antibody hybrids.

In addition, the antibodies offer the possibility of being tailored to many effector functions by selection in the immune system. This natural *protein engineering* system is unequalled in its efficiency. Such effector functions can be introduced into cells through the cytoplasmic expression of special functional antibody domains. Applications which lead to the modulation of cellular protein activity are also advantageous. This can occur, for example, by the stabilization of the target protein by protein-antibody complex formation. A change in the degradation kinetics can be achieved by this. Allosteric effector effects are also possible. The approximation of two effectors by the synthesis and stabilization of a ternary complex creates another possibility of influencing metabolic pathways, for example, by artificial multi-enzyme complexes or a local increase of metabolite concentrations at inducible operators. However, selecting the cytoplasmic expression of catalytic antibodies and the associated possibility of catalytic efficiency are especially advantageous. Cytoplasmic expression of functional antibodies is possible in a simple way for AB stabilized according to the invention.

Another advantage of the novel process is that antibodies which are formed inactive after expression under reducing conditions (e.g., DTE, DTT, glutathione) can be activated.

The antibodies stabilized according to the invention can be used advantageously in all fields of antibody application, for example, in the treatment of cancer and infections, as immunotoxins, for drug targeting, and in gene therapy. Use for imaging and in diagnostic procedures, for example, for the analysis of antigen-binding substances, is also advantageous.

The novel invention is especially advantageous for the stabilization of antibodies, which are already modified for other reasons, such as humanized or chimeric antibodies. This modification of amino acids can result in a destabilization, which leads to the restoration or even improvement

of the original stability of the antibodies by the novel process with additional modification of these antibodies outside the CDR regions.

Method for Analyzing Sequence Data Banks and Generating Tables 1-6
(canonical sequence approximation)

The invention assumes that the naturally occurring immunoglobulin sequences are a canonical collection of sequences, which in its sum should be compatible for all aspects of antibody functions.

The amino acid observed most frequently in a position in nature should be the one that best stabilizes a protein; this applies particularly to proteins in which stability and not special functions are naturally selected for.

At positions

35, 37, 39, 44, 45, 47, 50, 91, 95, 100j, 100k, 101, and 103

of the human heavy chain

and

1, 32, 34, 36, 38, 43, 44, 46, 49, 50, 87, 89, 95, 96, and 98

of the human light chain, the amino acids are involved in important interactions during the formation of heterodimeric Fv fragments; here stability is not primarily selected for. If the goal is improvement of dimerization properties, the most frequent amino acids at these positions are also to be selected; if the goal is improvement of stability, the second or third most frequent amino acids can also be selected.

The natural frequencies of the amino acids are determined for a random sample from the immunoglobulin data bank (Kabat, 1991). It is important that this sample represents the actual relations well. For this reason, it is also apparent that Tables 1 to 6 can be slightly modified if necessary if additional data become available. According to theory, the distribution within a species (e.g., human or mouse) or within a subtype (e.g., kappa or lambda) can be predicted.

Some closely related sequences are overrepresented in the data bank for procedural reasons. The result is that the data bank is not a suitable random sample without further modification. For this reason, only the largely complete sequences are selected from the data bank to avoid problems with the definition of a sequence distance between sequences which are known only as fragments. Sequences are selected for which more than 75% of the positions are known; this corresponds to not more than 30 missing positions for the light chains and not more than 33 missing positions in the heavy chains. Thus, the following are used as sequences from the Kabat data bank in the further analysis:

Protein	Number
V _L -kappa, mouse	731 of 1068 sequences
V _L -kappa, human	127 of 319 sequences
V _L -lambda, human	82 of 148 sequences
V _L -lambda, mouse	63 of 74 sequences
V _H , human	178 of 325 sequences
V _H , mouse	883 of 1620 sequences

In these sequences, all paired sequence distances are calculated and analyzed. Typically a two-peak distribution results with a maximum at the average distance for all sequences of this subtype. The distribution of the sequence distances can be used to reduce the effect of sampling errors: If the sequences of the natural distribution have a certain average distance and a certain distribution of distances, sampling errors are reduced, if a sample from the data bank is taken under the same limiting conditions of distance distribution. The limiting conditions employed are:

Protein	Minimum distance	Maximum distance
V _L -kappa, mouse	25	57
V _L -kappa, human	25	57
V _L -lambda, human	33	65

V _L -lambda, mouse	8	26
V _H , human	37	77
V _H , mouse	37	77

For this purpose, sequences are randomly selected from the data bank and checked to see if they satisfy a minimum distance and maximum distance to the previously selected sequences. If this is the case, they are assigned to the new sample. This is repeated until all sequences are checked for their suitability as a component of the sample. Typically, between 5 and 20% of the sequences in the data bank are selected. If this selection is repeated frequently (here 500 times), each individual sequence will be represented in the sample, nevertheless with a different frequency corresponding to its distance to the other sequences. Finally, the amino acid frequencies for the individual positions are determined from this new sample.

The *resampling* process described above was used for the frequencies determined here, with amino acids with a frequency below 0.1 (= 10%) not being listed in the tables.

The following examples, figures, tables, and the sequencing protocol describe the invention more closely:

Description of the Figures

Fig. 1. Expression plasmid pV_LH₃ (lac^{P_o}: promoter/operator region of the lac gene; ompA V_LH₃ coding region for the V_L domain with the signal sequence for the *outer membrane protein A*; t_{pp} terminator; fl-IG: F1 origin of replication of the phage; bla: b-lactamase gene; ori: origin of replication of the plasmid; lacI: lac-repressor gene). The drawing is not to scale.

(The plasmid corresponds substantially to the plasmid described in EP-B 0 324 162. The expression of antibodies is also described there.)

Fig. 2. Excitation spectrum of the V_L protein. Emission was measured at $I_{em} = 360$ nm. Protein concentration: 2 mM in PBS. The intensity is given in arbitrary units.

Fig. 3. Fluorescence spectrum of the folded (1) and unfolded (3) V_L protein and the difference spectrum (2). Excitation wavelength $I_{ex} = 284$ nm. Protein concentration: 2 mM in PBS. The intensity is given in arbitrary units.

Examples

Bacteria and Phages

E. coli K12 strains

CJ 236	dut1, ung1, thi-1, relA1 [pCJ 105 (Cam ^r), F'] (Geisselsoder et al., 1987) from Bio-Rad Laboratories GmbH, Munich
JM 83	ara, D(lac-pro AB) strA, thi-1 [F80lacZM15] (Yanisch-Perron et al., 1985)
JM 109	recA1, supE44, endA1, hsd R17, gyrA96, relA1, thi ₋ (lac-proAB) (Yanisch-Perron et al., 1985)

Bacteriophages

M13K07	Helper phage (Vieira & Messing, 1987) from Deutsche Pharmacia GmbH, Freiburg
--------	---

Plasmids

The plasmid p V_L H₅ codes for the V_L domain of the antibody McPC603 under the control of the lac promoter/operator. To purify the protein to homogeneity in one step by chromatography on immobilized zinc ions, the two C-terminal amino acids arginine (108) and alanine (109) are replaced by five histidine residues. In contrast to the wild type sequence, a leucine residue is in position 106 instead of an isoleucine.

For secretion into the periplasm, the signal sequence of the *outer membrane protein A* is inserted before the V_L coding region (Skerra & Plückthun, 1989).

Oligodeoxynucleotides

The employed oligodeoxynucleotides were prepared by the phosphoramidite process in a DNA synthesizer 380A from Applied Biosystems GmbH, Weiterstadt.

Growing of *E. coli* Cultures (Maniatis et al., 1982)

After incubation at 37°C with agitation at 180 to 200 rpm for 14 to 20 h, a dense overnight culture is obtained. Cell density is determined by measuring the OD_{600} . The medium is combined with a suitable antibiotic to select plasmid-bearing strains.

LB medium: 10 g/L of Bacto-Tryptone
5 g/L of Bacto-Yeast Extract
5 g/L of NaCl
2.5 mL/L of 1 M NaOH

Transformation of *E. coli* with Plasmid DNA (Hanahan, 1983)

Competent Cells

E. coli cells are made competent for transformation.

In a 250-mL Erlenmeyer flask, 20 mL of TYM medium was inoculated with 0.5 mL of a stationary overnight culture of the employed *E. coli* strain and incubated at 37°C to an OD₆₀₀ of 0.2 to 0.8. This culture is added to 100 mL of TYM medium. After growth to an OD₆₀₀ of 0.5 to 0.9, TYM medium is added to a total volume of 500 mL and the mixture is incubated further. When an OD₆₀₀ of 0.6 is reached, the bacterial suspension is cooled.

The bacteria are centrifuged at 4200 rpm and 4°C for 15 min; the supernatant is poured off, and the pellets are resuspended in 80 mL of Tfb I buffer and centrifuged; the supernatant poured off, and the pellets resuspended in a total of 16 mL of ice-cold Tfb II buffer.

TYM: 20 g/L of Bacto-Tryptone
 5 g/L of Bacto-Yeast Extract
 100 mM NaCl
 10 mM MgSO₄

<u>Tfb I:</u>	30 mM KOAc	<u>Tfb II:</u>	75 mM CaCl ₂
	50 mM MnCl ₂		10 mM KCl
	100 mM KCl		10 mM NaMOPS pH 7.0
	10 mM CaCl ₂		15% (v/v) glycerol
	15% (v/v) glycerol		

Transformation

The plasmid DNA is added to a volume of 30 µL of water and mixed well with the bacterial suspension. After 60 min on ice, the suspension is heat-shocked for 115 s at 37°C. After a

minute on ice, 800 μ L of LB medium is added; the bacterial suspension is transferred to a 10-mL culture tube and incubated for 60 min at about 180 rpm and 37°C. The total transformation batch is poured onto an LB plate with antibiotic and incubated for 12 to 16 h at 37°C.

Mutagenesis

Mutagenesis was carried out with use of the buffer from the Muta-Gene™ *in vitro* Mutagenesis Kit (Bio-Rad Laboratories GmbH, Germany) according to Kunkel (1985), Geisseloder et al., (1987), and Vieira and Messing (1987).

Production of Double Mutations by Recloning DNA Fragments

During the determination of the conformity stability of the individual mutants (2.3), Ala15Leu, Asn90Gln, and Phe32Tyr, inter alia, proved to be stabilizing. To analyze the additivity of the stabilizing effects, the double mutants Ala15Leu/Asn90Gln and Ala15Leu/Phe32Tyr were prepared.

The double mutants were prepared by recloning DNA fragments of already produced single mutants. After restriction digestion, the fragments were separated by agarose gel electrophoresis, and the desired fragments were cut out of the agarose; the DNA was isolated from these and ligated in a suitable manner.

Ala15Leu/Phe32Tyr:

The digestion of the plasmid DNA of the mutants Ala15Leu and Phe32Tyr using the restriction endonuclease Bst EII produced two fragments. One 3232-bp fragment contains the bases of amino acid 32, and a fragment of 870 bp those for amino acid 15. The difference in free enthalpy for unfolding into the unmodified antibody was found to be 22.6 kJ/mol (theoretically 20.8).

Ala15Leu/Asn90Gln:

The digestion of the plasmid DNA of the mutants Ala15Leu and Asn90Gln using the restriction endonuclease Xmn I produced two fragments. One 2991-bp fragment contained the bases for

amino acid 15, and a fragment of 1110 bp those for amino acid 90. The difference in free enthalpy for unfolding into the unmodified antibody was found to be 23.9 kJ/mol (theoretically 23.6).

Expression of Recombinant V_L Domains and Workup

Expression (Skerra and Plückthun, 1988) of the V_L proteins occurs under the control of the lac operator/repressor in *E. coli*, with the use of IPTG to induce expression. The signal sequence for the *outer membrane protein A* (ompA), which effects the secretion of the protein into the periplasm and is cleaved by *E. coli*'s own signal peptidase, is inserted before the protein coding region. Secretion into the periplasm because of the higher (oxidative) redox potential predominating there enables the formation of the central disulfide bonds and thus the correct folding of the V_L protein, which is not possible in the cytoplasm because of the lower (reducing) redox potential (Gilbert, 1990).

By selective lysis of the periplasm (1 M NaCl / 1 mM EDTA / 50 mM Tris/HCl, pH 8.0), the protein can be comfortably isolated in a mixture with other periplasmic proteins. The five C-terminal histidine residues, which are present instead of amino acids 108 and 109, enable simple purification of the protein to homogeneity in one step by chromatography on immobilized zinc ions (Hochuli et al., 1988).

LB medium (10 L) is inoculated with 200 mL of a stationary overnight culture of *E. coli* JM 83 / p V_LH_3 and combined with 10 mL of AMP stock solution. The culture is aerated and incubated at room temperature to OD₆₀₀ of 0.7 (about 4 h).

To induce V_LH_3 expression under the control of the lac operator/repressor, 5 mL of a 1 M IPTG solution is added, as well as 5 mL of the AMP stock solution which compensates for the loss of the selection antibiotic; the loss arises because the lysed bacteria release β -lactamase from the periplasm into the medium.

Incubation continues for another 3 h. To harvest the bacteria, the culture is filled to about 430 mL in a 500-mL centrifuge tube for rotor JA-10 of a Beckman centrifuge and centrifuged at 6000 rpm for 10 min in each case. Four centrifuge runs are necessary with six centrifuge tubes.

After the supernatant is poured off, typically about 30 g of bacterial pellets is obtained.

Periplasm Lysis

Two mL of periplasm lysis buffer is added per gram of cells; the bacteria are resuspended at 4°C with stirring and stirred further vigorously for at least an hour. Next, the milky pale-brown suspension is transferred to centrifuge tubes for rotor JA-20 of a Beckman centrifuge and the spheroblasts [sic] are separated by 20 min of centrifugation at 20,000 rpm and 4°C. The clear, pale-yellow supernatant with the recombinant V_L protein is transferred into a 50-mL Falcon tube and stored at 4°C until further use.

Chromatography on Immobilized Zinc Ions (Hochuli et al., 1988; Lindner et al., 1992)

The five C-terminal histidine residues of the V_L domain increase the binding of the protein to immobilized zinc ions so greatly that it can be purified in one step to homogeneity. The zinc is complexed to an iminodiacetate chelate ligand, which in turn is coupled to Sepharose. The histidine residues of the protein now function as complex ligands on zinc and are thus bound to the column material. Imidazole can be used for elution, which displaces the histidine residues on the zinc.

Preparation of the Column

To regenerate the column (about 5 mL of chelating Sepharose Fast Flow from Deutsche Pharmacia GmbH, Freiburg), it is first rinsed with 50 mL of the regeneration buffer and then with 20 mL of water to remove the complexed zinc and thus perhaps still bound protein. The column is then rinsed with 15 mL of zinc chloride solution (1 mg/mL), 15 mL of water, and 50 mL of the column equilibration buffer.

Chromatography

The chromatography is run at a flow rate of about 0.7 to 1 mL/min, with fractions being collected every 10 min.

After the application of the periplasm lysis [product] (typically about 70 mL), the column is rinsed with the column equilibration buffer until the OD₂₈₀ has returned to zero. Weakly bound proteins are eluted by rinsing with 10 mM imidazole in the column equilibration buffer. The elution of the V_LH₂ domain occurs in a linear gradient of 10 to 300 mM imidazole in the column equilibration buffer and a total volume of 200 mL in about 70 mM imidazole. The purity of the protein is verified by SDS-polyacrylamide gel electrophoresis.

Periplasm lysis buffer: 1 M NaCl
 1 mM EDTA
 50 mM Tris/HCl pH 8.0

Column equilibration buffer: 1 M NaCl
 50 mM Tris/HCl pH 8.0

Regeneration buffer: 1 M NaCl
 50 mM EDTA
 50 mM Tris/HCl pH 8.0

Next, the protein solution, which contains the desired amount of V_L protein (about 1 to 2 mg), is dialyzed twice against a hundredfold volume of the suitable buffer.

Determination of the Denaturation Curves

To determine the denaturation curves, the V_L protein is dialyzed against PBS and adjusted to a concentration of 0.2 mM (2.5 mg/mL; M = 12.4 kDa). To remove precipitates and other particles from the protein solution, said solution is centrifuged before use for 10 min in a refrigerated

centrifuge Sigma 2K15 and the supernatant is transferred to a new 1.5-mL Eppendorf reaction tube. Each 5 μ L of this protein solution is placed in a 5-mL test tube with a 10- μ L Hamilton syringe and combined with 500 μ L of denaturation buffer; the test tube is sealed with a silicone stopper and incubated overnight at 20°C.

Guanidinium chloride solutions in PBS in the concentration range of 0 to 5 M are used as denaturation buffer. Up to 2 M, the concentration is increased in 0.1-M increments, and above that in 0.2-M increments.

Equipment parameters:

Excitation wavelength: $I_{Ex} = 284$ nm

Emission wavelength: $I_{Em} = 360$ nm

Excitation slot width: 2 nm

Emission slot width: 10 nm

2.3. The Analysis of Denaturation Curves for Determining the Free Enthalpy of Unfolding

In the presence of denaturing compounds, proteins lose their native conformation and thus their biological function. Urea and guanidinium chloride are especially effective here. Many soluble globular proteins can be reversibly unfolded by these and exhibit a simple two-state behavior. This can be demonstrated by a comparison of calorimetric data (DH_{cal}) with the corresponding Van't Hoff enthalpies ($DH_{Van't Hoff}$), which can be determined from the temperature dependence of the equilibrium constants. The ratio of the two should be 1. It was demonstrated that deviations from this are very low for many single-domain proteins. This shows that possibly arising intermediates are thermodynamically unstable. They can be disregarded for this reason and the denaturation accordingly can be regarded as a cooperative transition between two macroscopic states, the folded (F) and unfolded (U) (Privalov, 1979).

The unfolded protein here is an ensemble of rapidly interconvertible conformers, which have the same or very similar energies. In the ideal case, the denatured state would form a random coil;

i.e., the rotation around bonds should be totally free and independent of all neighbors. Because the interactions of the solvent, the main chain atoms of the protein, and the 20 different side chains of amino acids cannot be ideal in the same way, a relation deviating from the ideal random coil is expected (Tanford, 1970). The spatial requirement of a real chain moreover will contribute to the retention of short-range interactions (Flory, 1969).

It can be assumed that in a concentrated guanidinium chloride solution a "complete" unfolding is achieved, which coincides with that achieved by other denaturing agents (Privalov, 1979; Creighton, 1978). However, the structure and dynamics of the individual groups in the denatured state can also be studied using the techniques of NMR spectroscopy (Wüthrich, 1986). This state appears as a high number of significantly different "polymorphic" conformers in rapid equilibrium (Dobson et al., 1985). The study of protein mutants provides strong evidence that the compactness of the denatured state, like its energy, can be greatly affected by individual mutations (Shortle, 1992).

The thermodynamic equilibrium constant K can be defined with the use of the two-state model:

$$F \rightleftharpoons U \quad (1)$$

$$K_u = [U]/[F] \quad (2)$$

The free enthalpy of the unfolding is obtained from this as:

$$\Delta G_u = -RT \ln K_u \quad (3)$$

The ratio $[U]/[F]$ can be determined by many spectroscopic measuring techniques, which detect a difference in the properties of the native and unfolded state, e.g., circular dichroism, UV absorption, or fluorescence spectroscopy. The last technique is regarded as the most sensitive and requires the smallest quantities of protein. A measured signal I consists additively of the folded (I_f) and unfolded (I_u) fractions:

$$I = I_u + I_f$$

These are proportional to the concentrations [F] and [U] present in equilibrium in each case.

With c_f and c_u as the proportionality constants, which are determined by the substance-specific, spectroscopic properties of both states,

$$I = c_f [F] + c_u [U] \quad (4)$$

is obtained.

With the mass balance ([P]: protein concentration)

$$[P] = [F] + [U], \quad (5)$$

by division of (4) and (5) after conversion with consideration of $c_f [P] = I_f^0$ and $c_u [P] = I_u^0$, which represent the signal intensities of the totally folded or unfolded states, respectively, we obtain:

$$\frac{[U]}{[F]} = \frac{I - I_f^0}{I_u^0 - I} \quad (6)$$

It can happen that the signal intensities of the completely folded or unfolded state depend on the concentration of the denaturing agent. This can be taken into account in good approximation by a linear correlation. In the case of the V_L domains in the present work, this applies to the unfolded state; it is taken into account with (7).

$$I_u^0 ([GdmHCl]) = I_u^0 - a \cdot [GdmHCl] \quad (7)$$

During the unfolding of a protein with a denaturing agent such as guanidium chloride, the stability of the protein is reduced with increasing denaturing agent concentration; in other words, ΔG_u becomes smaller. In the analysis of denaturing curves for proteins with a two-state behavior, it is assumed that there is a linear correlation between the concentration of the denaturing agent and ΔG_u (Pace, 1986; Alonso and Dill, 1991)

$$\Delta G_u = \Delta G_u^0 - m \cdot [D] \quad (8)$$

ΔG_u in the concentration range of the denaturing agent, in which both the folded and unfolded forms are present in detectable concentrations, can be calculated with use of (3) and (6). The free enthalpy of the unfolding in the absence of the denaturing agent is obtained then by linear extrapolation to zero molar denaturing agent, whereby the validity of (8) is used as a basis.

A second evaluation option is to derive an expression (9) for the signal intensity as a function of parameters with the use of (2), (3), (6), (7), and (8) and to obtain this by fitting the theoretical curve to the measured value by the principle of the least error squares.

$$I = I_u^0 - a \cdot [\text{GdmHCl}] + \frac{I_f^0 - I_u^0 + a \cdot [\text{GdmHCl}]}{1 + e^{\frac{m[\text{GdmHCl}] - \Delta G_u^0}{RT}}} \quad (9)$$

The following values occur as parameters: I_u^0 , I_f^0 , ΔG_u^0 , a , and m .

Fluorescence is used as a measurable variable to plot the denaturation curves for the V_L mutants. The fluorescence of the V_L protein is based primarily on the sole tryptophan residue 35, which is packed in the protein's interior against the central disulfide bond. During the unfolding, the tryptophan residue enters a more hydrophilic milieu and the interaction with the disulfide bond is lost. The very low fluorescence of the folded protein can be attributed to the fluorescence-quenching effect of the disulfide bond (Cowgill, 1967).

Figure 2 shows the fluorescence spectra of the folded and unfolded state (2 mM protein in PBS with 0 M or 5 M GdmHCl 20°C) and the difference spectrum for both. Over the course of unfolding, the protein fluorescence increases approximately by a factor of 16 with a maximum at 350 nm. The fluorescence in the present case thus proves to be an ideal measurable variable, because it undergoes a marked change with the unfolding of the protein. Figure 3 shows the excitation spectrum; the fluorescence at 350 [nm] is determined as function of the excitation wavelength. A pronounced maximum at 280 nm is evident.

PBS 4 mM KH_2PO_4
 16 mM Na_2HPO_4
 115 mM NaCl
 The pH is 7.4.

Several measurements were generally performed; the data were obtained by averaging the values standardized using (10) (from (5), (6), and (7)).

$$\frac{[U]}{[P]} = \frac{1 - I_i^0}{I_u^0 - a \cdot [\text{GdmHCl}] - I_i^0} \quad (10)$$

The concentration at which half of the proteins are unfolded, the denaturation midpoint $[\text{GdmHCl}]_{1/2}$, can be calculated from the obtained parameters. $\text{DG}_u = 0$ applies here. (11) is obtained from (8).

$$[\text{GdmHCl}]_{1/2} = \frac{\Delta G_u^0}{m} \quad (11)$$

The parameters from the individual measurements are summarized in Table 7.

Table 7. Comparison between Prediction and Experiment for Stabilizing Point Mutations

Domain	f_{WT}	f_{mut}	ΔG^{Pfold} (kJ mol ⁻¹)	Experiment	Prediction
WT			-13.5		
Ala15Leu	0.082	0.411	-19.2	++	++
Asn90Gln	0.047	0.892	-17.9	++	++
Phe32Tyr	0.034	0.799	-15.1	+	++
Leu106Ile	0.298	0.684	-15.0	+	+
Thr63Ser	0.148	0.823	-14.7	+	++
Met21Ile	0.278	0.590	-14.5	+	+
Met21Leu	0.278	0.103	-12.2	-	-

During the expression of the proteins, it was surprisingly observed in addition that the yield in comparison with the wild type rose to about 9 to 26 mg in more stable mutants.

7. Bibliography

- Alonso, D. O. V., Dill, K. A. (1991). Solvent Denaturation and Stabilization of Globular Proteins. *Biochemistry* 30, 5974-5985
- Birnboim, C. Doly, J. (1979). Rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids. Res.* 7, 1513-1523
- Cowgill, R. W. (1967). Fluorescence Quenching by Disulfide and Sulfhydryl Groups. *Biochimica et Biophysica acta* 140, 37-44
- Creighton, T. E. (1978). Experimental Studies of Protein Folding and Unfolding. *Prog. Biophys. molec. Biol.* 33, 231-297
- Devereux et al., *Nucleic Acids Res.* 12 (1984), 387 - 395
- Dobson, C. M. Evans, P. A., Fox, R. O. (1985). In: *Structure and Motion. Membranes Nucleic Acids and Proteins*, Adenine, Gunderland, New York, 265-276
- Flory, P. (1969). *Statistical Mechanics of Chain Molecules*, Wiley, New York, 432 ff.
- Frisch C. et al., *Biol. Chem. Hoppe-Seyler* 375 (1994) 353 - 356
- Geisselsoder, J., Witney, F., Yuckenberg, P. (1987). Efficient site-directed *in vitro* mutagenesis. *Biotechniques* 5, 786-791
- Gellert, W. (Hsg.) (1984). *Kleine Enzyklopädie Mathematik*, 2. Auflage, Verlag Harri Deutsch, Thun Frankfurt/M., 668 ff.
- Gilbert, H. F. (1990). Molecular and cellular aspects of thiol-disulfide exchange. *Adv. Enzymol.* 63, 69-172
- Glockshuber, R. (1989). Das F_V-Fragment des Phosphorylcholin bindenden Antikörpers McPC603: Expression in *Escherichia coli* und Charakterisierung. *Dissertation*, Fakultät für Chemie und Pharmazie der Ludwig-Maximilians-Universität, München
- Hanahan, D. (1983). Studies of transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166, 557-579

- Hochuli, E., Bannworth, W., Döbeli, H., Gentz, R., Süber, D. (1988). Genetic approach to facilitate purification of recombinant proteins with a novel metal chelat adsorbent. *Bio/Technology* 6, 1321-1325
- Kaplan, B. E. (1985). The automated synthesis of oligodeoxyribonucleotides. *Trends Biotechnol.* 3, 253-256
- Kunkel, T. A. Roberts, J. D., Zakow, R. A. (1987). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.* 154, 367-382
- Maniatis, T. Fritsch, E. F., Sambrook, J. (1982). *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory, New York
- Morrison, S.L. et al., *Ann. Rev. Immunol.* 2 (1984) 239 - 256
- Pace, C. N. (1986). Determination and Analysis of Urea and Guanidine Hydrochloride Denaturation Curves. *Methods Enzymol.* 131, 266
- Privalov, P. L., Gill, S. J. (1989). The hydrophobic effect: a reappraisal. *Pure & Appl. Chem.* 61, 1097-1104
- Shortle, D. (1992). Mutational studies of protein structures and their stabilities. *Q. Rev. Bioph.* 25, 205-250
- Skerra, A., Plückthun, A. (1988). Assembly of functional F_v fragments in *Escherichia coli*. *Science* 240, 1038-1041
- Tanford, C. (1970). Protein Denaturation. Part C. Theoretical Models for the Mechanism of Denaturation. *Adv. Protein Chem.* 24, 1-95
- Vieira, J. Messing, J. (1987). Production of single-stranded plasmid DNA. *Methods Enzymol.* 153, 3-11
- Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids*, Wiley, New York, 293 ff.
- Yanisch-Perron, C., Vieira, J. Messing, J. (1985). Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33, 103-119
- Frisch, C. et. al., *Biol. Chem. Hoppe-Seyler* 375 375 (1994) 353 - 356
- Rollence, M.L., Filpula, D., Pantoliano, M.W. & Bryan P.N. *Crit. Rev. Biotechnol.* 8, 217 - 224 (1988)
- Chen, L.H. & Baldwin, T.O. *Biochemistry* 28, 2684 - 2689 (1989)
- Turner, S.L., Ford, G.C., Mountain, A. & Moir, A. *Prot. Eng.* 5, 535 - 541 (1992)
- Risse, B., Stempfer, G., Rudolph, R., Schumacher, G., & Jaenicke, R. *Protein Science* 1, 1710 - 1718 (1992).

surprisingly observed in addition that the yield in
9 to 26 mg in more stable mutants.

ent Denaturation and Stabilization of
y 30, 5974-5985

ne extraction procedure for screening
Acids. Res. 7, 1513-1523

aching by Disulfide and Sulfhydryl

r 140, 37-44

udies of Protein Folding and Unfolding.

3, 231-297

984), 387 - 395

1985). In: *Structure and Motion*.

denine, Guilderland, New York, 265-

Chain Molecules, Wiley, New York,

r 375 (1994) 353 - 356

P. (1987). Efficient site-directed *in vitro*
, 786-791

opädie Mathematik, 2. Auflage, Verlag
t/M., 668 ff.

lar aspects of thiol-disulfide exchange.

t des Phosphorylcholin bindenden

ssion in *Escherichia coli* und

n, Fakultät für Chemie und Pharmazie

ersität, München

iation of *Escherichia coli* with plasmids.

Hochuli, E., Barinworth, W., Dö
approach to facilitate

metal chelat adsorber

Kaplan, B. E. (1985). The autom
Trends Biotechnol. 3

Kunkel, T. A. Roberts, J. D., Za
specific mutagenesis
154, 367-382

Maniatis, T. Fritsch, E. F., Samb
manual. Cold Spring

Morrison, S.L. et al., *Ann. Rev.*

Pace, C. N. (1986). Determinatio
Hydrochloride Denat

Privalov, P. L., Gill, S. J. (1989)
Appl. Chem. 61, 109

Shortle, D. (1992). Mutational s
Q. Rev. Bioph. 25, 2

Skerra, A., Plückthun, A. (1988)
Escherichia coli. Sc

Tanford, C. (1970). Protein Den
Mechanism of Denat

Vieira, J. Messing, J. (1987). Pr
Methods. Enzymol.

Wüthrich, K. (1986). *NMR of P*
293 ff.

Yanisch-Perron, C., Vieira, J. M.
vectors and host str

pUC19 vectors. *Ger*

Frisch, C. et. al., *Biol. Chem. H*
Rollence, M.L., Filpula, D., Par

Biotechnol. 8, 21

Chen, L.H. & Baldwin, T.O. Bi

Turner, S.L., Ford, G.C., Mour

Risse, B., Stempfer, G., Rudolf

Science 1, 1710 - 1

- Hochuli, E., Bannworth, W., Döbeli, H., Gentz, R., Süber, D. (1988). Genetic approach to facilitate purification of recombinant proteins with a novel metal chelat adsorbent. *Bio/Technology* 6, 1321-1325
- Kaplan, B. E. (1985). The automated synthesis of oligodeoxyribonucleotides. *Trends Biotechnol.* 3, 253-256
- Kunkel, T. A. Roberts, J. D., Zakow, R. A. (1987). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.* 154, 367-382
- Maniatis, T. Fritsch, E. F., Sambrook, J. (1982). *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory, New York
- Morrison, S.L. et al., *Ann. Rev. Immunol.* 2 (1984) 239 - 256
- Pace, C. N. (1986). Determination and Analysis of Urea and Guanidine Hydrochloride Denaturation Curves. *Methods Enzymol.* 131, 266
- Privalov, P. L., Gill, S. J. (1989). The hydrophobic effect: a reappraisal. *Pure & Appl. Chem.* 61, 1097-1104
- Shortle, D. (1992). Mutational studies of protein structures and their stabilities. *Q. Rev. Bioph.* 25, 205-250
- Skerra, A., Plückthun, A. (1988). Assembly of functional Fv fragments in *Escherichia coli*. *Science* 240, 1038-1041
- Tanford, C. (1970). Protein Denaturation. Part C. Theoretical Models for the Mechanism of Denaturation. *Adv. Protein Chem.* 24, 1-95
- Vieira, J. Messing, J. (1987). Production of single-stranded plasmid DNA. *Methods. Enzymol.* 153, 3-11
- Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids*, Wiley, New York, 293 ff.
- Yanisch-Perron, C., Vieira, J. Messing, J. (1985). Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33, 103-119
- Frisch, C. et. al., *Biol. Chem. Hoppe-Seyler* 375 375 (1994) 353 - 356
- Rollence, M.L., Filpula, D., Pantoliano, M.W. & Bryan P.N. *Crit. Rev. Biotechnol.* 8, 217 - 224 (1988)
- Chen, L.H. & Baldwin, T.O. *Biochemistry* 28, 2684 - 2689 (1989)
- Turner, S.L., Ford, G.C., Mountain, A. & Moir, A. *Prot. Eng.* 5, 535 - 541 (1992)
- Risse, B., Stempffer, G., Rudolph, R., Schumacher, G., & Jaenicke, R. *Protein Science* 1, 1710 - 1718 (1992).

- Arase, A., Yomo, T., Urabe, I., Hata, Y., Katsube, Y. & Okada, H. FEBS lett., 316, 123 - 127 (1993).
- Kabat, E.A., Wu, T.T., Perry, H.M., Gottesman, K.S. & Foeller, C. Sequences of Proteins of Immunological Interest, 5th ed. US Dept. Health and Human Services, Bethesda, Md. (1991)
- Tonegawa, S. Nature 302, 575 - 581 (1983)
- Berek, C. & Milstein, C., Immunol. Rev. 105, 5 - 26 (1988)
- French, D.L., Laskov, R. & Scharff, M.D. Science, 244, 1152 - 1157 (1989)
- Berek, C. & Ziegner, M. Immunol. Today 14, 400 - 404 (1993)
- Jung, S.H. et al., Proteins: Structure Function and Genetics 19 (1994) 35 - 47
- Glockshuber, R. et al., Biochemistry 29 (1990) 1362 - 1367
- Lindner, P. et al., Methods: A comparison to Methods of Enzymology 4 (1992) 41 - 56

Table 1

Variable domains of the human heavy chain

Determined consensus sequence

Table 1

```
1  EVQLVESGGGLVKPGGSLRLSCAASGFTFSSYAMS--WVRQA
41 PGKGLEWVGWIY---NGGDTYYADSVKGRFTISRDTSKNTLYL
81 QMNSLRAEDTAVYYCARGGGGGY-----FDYWGQGTLLTVSS
```

(SEQ ID NO:1 shows the consensus sequence without introns)

Position Type and Frequency

```
_ 1 :  E:0.505 Q:0.460
_ 2 :  V:0.939
_ 3 :  Q:0.920
_ 4 :  L:0.997
_ 5 :  V:0.680 Q:0.177 L:0.106
_ 6 :  E:0.638 Q:0.333
_ 7 :  S:0.930
_ 8 :  G:1.000
_ 9 :  G:0.566 A:0.215 P:0.177
_10 :  G:0.609 E:0.246
_11 :  L:0.637 V:0.336
_12 :  V:0.701 K:0.221
_13 :  K:0.455 Q:0.455
_14 :  P:0.964
_15 :  G:0.796 S:0.177
_16 :  G:0.357 R:0.189 Q:0.124 E:0.116 A:0.116
_17 :  S:0.803 T:0.179
_18 :  L:0.794 V:0.188
```

- 30 -

_19: R:0.574 K:0.230 S:0.164
_20: L:0.730 V:0.217
_21: S:0.786 T:0.214
_22: C:1.000
_23: A:0.506 K:0.228 T:0.132
_24: A:0.655 V:0.174
_25: S:0.977
_26: G:0.985
_27: F:0.578 Y:0.161 G:0.147
_28: T:0.585 S:0.266
_29: F:0.804
_30: S:0.761 T:0.110
_31: S:0.389 D:0.167 T:0.160
_32: Y:0.535 S:0.124
_33: A:0.292 Y:0.168 W:0.127
_34: M:0.598 I:0.183
_35: S:0.310 H:0.272 N:0.112
_35a: -:0.897
_35b: -:0.922
_36: W:1.000
_37: V:0.784 I:0.151
_38: R:1.000
_39: Q:0.994
_40: A:0.648 P:0.162
_41: P:0.923
_42: G:0.969
_43: K:0.729 Q:0.156 R:0.104
_44: G:0.909
_45: L:0.959
_46: E:0.972
_47: W:0.996
_48: V:0.566 M:0.196 I:0.150
_49: G:0.510 S:0.243 A:0.197
_50: W:0.199 V:0.113
_51: I:0.807
_52: Y:0.211 S:0.167 N:0.115 G:0.107

- 31 -

_ 52a: -:0.198 P:0.159 Y:0.128 G:0.113
_ 52b: -:0.897
_ 52c: -:0.927
_ 53 : N:0.170 D:0.166 S:0.134 G:0.125
_ 54 : G:0.402 S:0.204
_ 55 : G:0.476 S:0.269
_ 56 : D:0.198 S:0.183 T:0.158 N:0.143
_ 57 : T:0.465 K:0.105
_ 58 : Y:0.304 N:0.186 H:0.114
_ 59 : Y:0.894
_ 60 : A:0.656 N:0.129
_ 61 : D:0.394 P:0.205 Q:0.138
_ 62 : S:0.714 K:0.122
_ 63 : V:0.590 F:0.219 L:0.154
_ 64 : K:0.554 Q:0.237
_ 65 : G:0.785 S:0.148
_ 66 : R:0.926
_ 67 : F:0.602 V:0.348
_ 68 : T:0.878
_ 69 : I:0.806 M:0.111
_ 70 : S:0.789 T:0.130
_ 71 : R:0.597 V:0.150
_ 72 : D:0.815 N:0.152
_ 73 : T:0.301 N:0.284 D:0.253
_ 74 : S:0.890 A:0.103
_ 75 : K:0.643
_ 76 : N:0.672 S:0.221
_ 77 : T:0.659 Q:0.211
_ 78 : L:0.462 A:0.252 F:0.179
_ 79 : Y:0.710 S:0.192
_ 80 : L:0.822 M:0.169
_ 81 : Q:0.573 E:0.198
_ 82 : M:0.548 L:0.344
_ 82a: N:0.399 S:0.300
_ 82b: S:0.797
_ 82c: L:0.753 V:0.202

- 32 -

_83 : R:0.542 T:0.196 K:0.131
_84 : A:0.485 P:0.191 S:0.134
_85 : E:0.644 A:0.155 D:0.127
_86 : D:0.972
_87 : T:0.940
_88 : A:0.956
_89 : V:0.765
_90 : Y:0.992
_91 : Y:0.947
_92 : C:0.998
_93 : A:0.891
_94 : R:0.681 K:0.158
_95 : G:0.179 D:0.152 E:0.119 V:0.100
_96 : G:0.118 P:0.101
_97 : G:0.168 S:0.122
_98 : G:0.132 Y:0.103
_99 : G:0.240 A:0.111
_100 : Y:0.139 S:0.127 -:0.127 G:0.120
_100a: -:0.276 S:0.160
_100b: -:0.379 S:0.107 G:0.101
_100c: -:0.429 Y:0.110
_100d: -:0.567
_100e: -:0.645 Y:0.129
_100f: -:0.728 Y:0.107
_100g: -:0.758 Y:0.114
_100h: -:0.825
_100i: -:0.868
_100j: -:0.481 Y:0.147
_100k: F:0.475 -:0.176 M:0.160 L:0.100
_101 : D:0.755
_102 : Y:0.442 V:0.239
_103 : W:0.967
_104 : G:0.953
_105 : Q:0.823
_106 : G:1.000
_107 : T:0.887

_108: L:0.659 T:0.194
_109: V:0.986
_110: T:0.916
_111: V:0.969
_112: S:0.980
_113: S:0.930

Table 2

Variable domains of the mouse heavy chain

Determined consensus sequence

```
1  EVQLQQSGGELVKPGASVKLSCKASGYTFTSYMH--WVKQR
41 PGKGLEWIGRINP--GSGGTNYNEKFKGKATLTRDKSSSTAYL
81 QLSSLTSEDSAVYYCARGGY-----YFDYWGQGTTVTVSS
```

(SEQ ID NO:2 shows the consensus sequence without introns)

Position Type and Frequency

```
_ 1 :  E:0.504 Q:0.409
_ 2 :  V:0.965
_ 3 :  Q:0.756 K:0.186
_ 4 :  L:0.968
_ 5 :  Q:0.575 V:0.227
_ 6 :  Q:0.563 E:0.434
_ 7 :  S:0.818 P:0.122
_ 8 :  G:0.976
_ 9 :  G:0.314 P:0.311 A:0.246 T:0.107
_10 :  E:0.560 G:0.353
_11 :  L:0.951
_12 :  V:0.810
_13 :  K:0.526 Q:0.248 R:0.118
_14 :  P:0.895
_15 :  G:0.883
_16 :  A:0.383 G:0.314
_17 :  S:0.940
_18 :  V:0.599 L:0.290
_19 :  K:0.738 S:0.100
_20 :  L:0.569 I:0.245 M:0.173
_21 :  S:0.915
```

- 35 -

_22 : C:1.000
_23 : K:0.528 A:0.222 T:0.121
_24 : A:0.779
_25 : S:0.912
_26 : G:0.988
_27 : Y:0.591 F:0.380
_28 : T:0.671 S:0.171
_29 : F:0.858
_30 : T:0.578 S:0.323
_31 : S:0.351 D:0.276 N:0.122
_32 : Y:0.723
_33 : Y:0.312 W:0.298 G:0.163
_34 : M:0.664 I:0.199
_35 : H:0.300 N:0.283 S:0.181
_35a: -:0.971
_35b: -:0.998
_36 : W:0.997
_37 : V:0.909
_38 : K:0.550 R:0.434
_39 : Q:0.945
_40 : R:0.384 S:0.170 A:0.143 T:0.105
_41 : P:0.866
_42 : G:0.750 E:0.195
_43 : K:0.525 Q:0.321
_44 : G:0.671 R:0.108 S:0.102
_45 : L:0.981
_46 : E:0.930
_47 : W:0.944
_48 : I:0.647 V:0.176 L:0.102
_49 : G:0.742 A:0.250
_50 : R:0.196 Y:0.157 E:0.103
_51 : I:0.921
_52 : N:0.295 Y:0.185 S:0.147 D:0.116 R:0.101
_52a: P:0.550 S:0.148
_52b: -:0.893 K:0.104
_52c: -:0.891

- 36 -

_ 53 : G:0.321 N:0.190 Y:0.162 S:0.102
_ 54 : S:0.310 N:0.309 G:0.222
_ 55 : G:0.568 S:0.153 Y:0.107
_ 56 : G:0.162 Y:0.158 S:0.149 T:0.126 N:0.117
_ 57 : T:0.763 I:0.115
_ 58 : N:0.295 Y:0.183 K:0.161
_ 59 : Y:0.956
_ 60 : N:0.536 A:0.181
_ 61 : E:0.294 Q:0.197 D:0.184 P:0.150
_ 62 : K:0.508 S:0.269
_ 63 : F:0.607 V:0.221 L:0.131
_ 64 : K:0.809
_ 65 : G:0.596 D:0.174 S:0.166
_ 66 : K:0.532 R:0.466
_ 67 : A:0.516 F:0.341
_ 68 : T:0.773
_ 69 : L:0.437 I:0.417
_ 70 : T:0.590 S:0.373
_ 71 : R:0.339 V:0.306 A:0.230
_ 72 : D:0.895
_ 73 : K:0.366 N:0.258 T:0.238
_ 74 : S:0.764 A:0.123
_ 75 : S:0.539 K:0.254
_ 76 : S:0.585 N:0.334
_ 77 : T:0.772
_ 78 : A:0.514 L:0.269 V:0.168
_ 79 : Y:0.868 F:0.108
_ 80 : L:0.481 M:0.475
_ 81 : Q:0.742 E:0.137
_ 82 : L:0.589 M:0.342
_ 82a: S:0.525 N:0.286
_ 82b: S:0.710 N:0.109
_ 82c: L:0.891 V:0.101
_ 83 : T:0.587 R:0.231 K:0.104
_ 84 : S:0.736
_ 85 : E:0.876

- 37 -

_86 : D:0.981
_87 : S:0.513 T:0.428
_88 : A:0.941
_89 : V:0.542 T:0.150 M:0.126 I:0.104
_90 : Y:0.980
_91 : Y:0.770 F:0.227
_92 : C:0.997
_93 : A:0.725 T:0.127
_94 : R:0.821
_95 : G:0.174 D:0.158 Y:0.125
_96 : G:0.205 Y:0.150
_97 : Y:0.242 G:0.181
_98 : Y:0.249 -:0.208 G:0.149
_99 : -:0.310 G:0.181 S:0.105
_100 : -:0.484 S:0.110 G:0.108
_100a: -:0.612
_100b: -:0.768
_100c: -:0.867
_100d: -:0.910
_100e: -:0.974
_100f: -:0.986
_100g: -:0.992
_100h: -:0.997
_100i: -:1.000
_100j: Y:0.324 -:0.276 A:0.177 W:0.101
_100k: F:0.590 M:0.187 -:0.108
_101 : D:0.668 A:0.233
_102 : Y:0.774 V:0.152
_103 : W:0.986
_104 : G:0.985
_105 : Q:0.834
_106 : G:0.993
_107 : T:0.973
_108 : T:0.475 L:0.283 S:0.226
_109 : V:0.695 L:0.294
_110 : T:0.960

- 38 -

_111 : V:0.987
_112 : S:0.984
_113 : S:0.753 A:0.233

ght chain

ASQSISS-----YLAWYQQKPGKAPKLLIYD
ISSLPEDFATYYCQYYSLP-----YTFGQ

quence without introns)

- 41 -

S:0.951
N:0.389 S:0.292 T:0.252
L:0.693 R:0.307
E:0.335 Q:0.219 A:0.168
S:0.490 T:0.333
G:0.981
V:0.788 I:0.183
P:0.992
S:0.745 D:0.160
R:0.944
F:0.989
S:0.889
G:0.992
S:0.860
G:0.882
S:0.980
G:0.968
T:0.954
D:0.782 E:0.180
F:0.977
T:0.898
L:0.766 F:0.223
T:0.930
I:0.953
S:0.870
S:0.669 R:0.136 G:0.131
L:0.907
Q:0.753 E:0.164
P:0.817
E:0.803 D:0.177
D:0.957
F:0.656 V:0.143 I:0.140
A:0.945
T:0.598 V:0.283
Y:0.989
Y:0.919

Table 4

Variable domains of the mouse kappa light chain

Determined consensus sequence

```
1 DIVMTQSPASLSASLGERVTITCRASQSVSS-----YLHWYQQKPGQSPKLLIYR
51 ASNLASGVDPDRFSGSGSGTDFTLTISSVEAEDLATYYCQQSNSYP-----YTFGG
101 GTKLEI-KR
```

(SEQ ID NO:4 shows the consensus sequence without introns)

Position Type and Frequency

```
_ 1: D:0.707 E:0.106 Q:0.104
_ 2: I:0.813 V:0.121
_ 3: V:0.653 Q:0.223
_ 4: M:0.520 L:0.424
_ 5: T:0.878
_ 6: Q:0.995
_ 7: S:0.808 T:0.152
_ 8: P:0.871
_ 9: A:0.383 S:0.313 K:0.115
_10: S:0.571 I:0.177
_11: L:0.578 M:0.322
_12: S:0.476 A:0.257 P:0.117
_13: A:0.482 V:0.393
_14: S:0.874
_15: L:0.464 P:0.273 V:0.133
_16: G:0.977
_17: E:0.462 D:0.313 Q:0.188
_18: R:0.447 K:0.282
_19: V:0.693 A:0.192 I:0.103
_20: T:0.708 S:0.253
_21: I:0.607 M:0.224 L:0.122
```

- 44 -

_22 : T:0.487 S:0.432
_23 : C:0.984
_24 : R:0.476 K:0.253 S:0.161
_25 : A:0.812 S:0.166
_26 : S:0.974
_27 : Q:0.524 S:0.231 E:0.133
_28 : S:0.623 D:0.157 N:0.118
_29 : V:0.411 I:0.383 L:0.176
_30 : S:0.408 G:0.129
_31 : S:0.227 N:0.192 T:0.158 -0.111
_31a: -0.569 S:0.256
_31b: -0.685 G:0.141
_31c: -0.685 G:0.102
_31d: -0.690 S:0.112
_31e: -0.821 T:0.103
_31f: -0.924
_32 : Y:0.652 N:0.122
_33 : L:0.603 M:0.231 V:0.114
_34 : H:0.330 A:0.227 N:0.155
_35 : W:0.989
_36 : Y:0.790 F:0.126
_37 : Q:0.893 L:0.102
_38 : Q:0.926
_39 : K:0.879
_40 : P:0.808 S:0.134
_41 : G:0.773
_42 : Q:0.450 K:0.151 G:0.105
_43 : S:0.641 P:0.154 T:0.114
_44 : P:0.888
_45 : K:0.810
_46 : L:0.802
_47 : L:0.814 W:0.171
_48 : I:0.938
_49 : Y:0.880
_50 : R:0.186 Y:0.164 S:0.147 G:0.137 A:0.102
_51 : A:0.507 T:0.296

- 45 -

_ 52 : S:0.893
_ 53 : N:0.469 T:0.146
_ 54 : L:0.630 R:0.262
_ 55 : A:0.322 E:0.182 Y:0.123
_ 56 : S:0.687 T:0.140
_ 57 : G:0.997
_ 58 : V:0.864 I:0.132
_ 59 : P:0.961
_ 60 : D:0.354 S:0.279 A:0.262
_ 61 : R:0.976
_ 62 : F:0.997
_ 63 : S:0.776 T:0.192
_ 64 : G:0.992
_ 65 : S:0.981
_ 66 : G:0.955
_ 67 : S:0.969
_ 68 : G:0.896
_ 69 : T:0.862
_ 70 : D:0.675 S:0.226
_ 71 : F:0.618 Y:0.373
_ 72 : T:0.554 S:0.434
_ 73 : L:0.901
_ 74 : T:0.702 K:0.111
_ 75 : I:0.977
_ 76 : S:0.716 N:0.107
_ 77 : S:0.567 P:0.142 R:0.127
_ 78 : V:0.483 L:0.268 M:0.232
_ 79 : E:0.667 Q:0.284
_ 80 : A:0.479 S:0.124 P:0.115 E:0.106
_ 81 : E:0.878 D:0.115
_ 82 : D:0.976
_ 83 : L:0.261 A:0.223 F:0.149 I:0.108 V:0.108
_ 84 : A:0.764 G:0.170
_ 85 : T:0.446 V:0.216 D:0.120
_ 86 : Y:0.995
_ 87 : Y:0.712 F:0.259

- 46 -

_88 : C:0.999
_89 : Q:0.669 L:0.131
_90 : Q:0.905
_91 : S:0.196 G:0.196 H:0.193 Y:0.117 W:0.100
_92 : N:0.224 S:0.223 Y:0.169
_93 : S:0.395 E:0.199
_94 : Y:0.285 L:0.114 S:0.101
_95 : P:0.938
_95a: -:0.957
_95b: -:0.990
_95c: -:1.000
_95d: -:1.000
_95e: -:1.000
_95f: -:1.000
_96 : Y:0.263 L:0.255 W:0.172 R:0.114
_97 : T:0.992
_98 : F:1.000
_99 : G:0.996
_100 : G:0.631 A:0.244 S:0.114
_101 : G:0.997
_102 : T:1.000
_103 : K:0.974
_104 : L:0.995
_105 : E:0.998
_106 : I:0.763
_106a: -:1.000
_107 : K:0.958
_108 : R:1.000

Table 5

Variable domains of the human lambda light chain

Determined consensus sequence

```
1 QSELTQPPS-VSVSPGQTVTISCSGDSLIG-----YVSWYQQKPGQAPKLVIIYD
51 DNKRPSGIPDRFSGSKSGNTASLTISGLQAEDEADYYCQSWDSSS-----VVFGG
101 GTKLTVLGQP
```

(SEQ ID NO:5 shows the consensus sequence without introns)

Position Type and Frequency

```
_ 1: Q:0.557 S:0.211
_ 2: S:0.486 Y:0.392
_ 3: E:0.299 A:0.271 V:0.239
_ 4: L:0.995
_ 5: T:0.920
_ 6: Q:1.000
_ 7: P:0.865
_ 8: P:0.704 A:0.126
_ 9: S:0.911
_10: -:1.000
_11: V:0.858
_12: S:0.974
_13: V:0.410 G:0.345 A:0.129
_14: S:0.656 A:0.259
_15: P:0.826 L:0.123
_16: G:0.960
_17: Q:0.837
_18: T:0.544 S:0.291 R:0.111
_19: V:0.434 A:0.391 I:0.126
_20: T:0.518 R:0.259
_21: I:0.888
```

_ 22 : S:0.518 T:0.457
_ 23 : C:1.000
_ 24 : S:0.471 T:0.243
_ 25 : G:0.903
_ 26 : D:0.389 S:0.214 T:0.183
_ 27 : S:0.380 N:0.123 T:0.100
_ 28 : L:0.366 S:0.322
_ 29 : G:0.225 D:0.221 N:0.194 P:0.117
_ 30 : I:0.264 V:0.230 K:0.102
_ 31 : G:0.303 K:0.151 A:0.129
_ 31a: -:0.449 S:0.133 G:0.114 D:0.113
_ 31b: -:0.486 N:0.168 Y:0.147
_ 31c: -:0.682 N:0.166
_ 31d: -:1.000
_ 31e: -:1.000
_ 31f: -:1.000
_ 32 : Y:0.413 S:0.211 F:0.104 H:0.100
_ 33 : V:0.647 A:0.228
_ 34 : S:0.429 H:0.126 Y:0.110
_ 35 : W:0.999
_ 36 : Y:0.856
_ 37 : Q:0.946
_ 38 : Q:0.867
_ 39 : K:0.275 R:0.229 H:0.215 L:0.132
_ 40 : P:0.921
_ 41 : G:0.846
_ 42 : Q:0.453 K:0.224 T:0.156
_ 43 : A:0.770 S:0.171
_ 44 : P:0.999
_ 45 : K:0.400 V:0.319 L:0.111
_ 46 : L:0.777
_ 47 : V:0.542 L:0.306 I:0.103
_ 48 : I:0.822 V:0.131
_ 49 : Y:0.824 F:0.123
_ 50 : D:0.284 E:0.254
_ 51 : D:0.338 V:0.194 N:0.173

- 49 -

_52 : N:0.386 S:0.260 T:0.191
_53 : K:0.255 Q:0.149 N:0.147 D:0.120
_54 : R:0.950
_55 : P:0.905
_56 : S:0.875
_57 : G:0.873
_58 : I:0.595 V:0.369
_59 : P:0.875 S:0.082
_60 : D:0.392 E:0.326 L:0.109
_61 : R:0.966
_62 : F:0.967
_63 : S:0.999
_64 : G:0.913
_65 : S:0.974
_66 : K:0.437 S:0.193 N:0.190
_67 : S:0.959
_68 : G:0.859
_69 : N:0.520 T:0.242
_70 : T:0.565 S:0.275
_71 : A:0.913
_72 : S:0.491 T:0.436
_73 : L:0.999
_74 : T:0.812 A:0.116
_75 : I:0.945
_76 : S:0.718 T:0.208
_77 : G:0.828 R:0.120
_78 : L:0.534 V:0.194 A:0.154 T:0.117
_79 : Q:0.656 E:0.165
_80 : A:0.460 S:0.175 T:0.171 V:0.127
_81 : E:0.573 G:0.174
_82 : D:0.971
_83 : E:0.993
_84 : A:0.974
_85 : D:0.908
_86 : Y:0.999
_87 : Y:0.817 F:0.183

- 50 -

_88 : C:0.999
_89 : Q:0.473 S:0.203
_90 : S:0.524 T:0.208 A:0.190
_91 : W:0.438 Y:0.336
_92 : D:0.590
_93 : S:0.388 N:0.160 D:0.148
_94 : S:0.537 G:0.156
_95 : S:0.245 G:0.167 L:0.158
_95a: -:0.343 S:0.156 N:0.103
_95b: -:0.590
_95c: -:0.941
_95d: -:0.992
_95e: -:1.000
_95f: -:1.000
_96 : V:0.344 P:0.101
_97 : V:0.715 I:0.152 L:0.111
_98 : F:0.999
_99 : G:0.999
_100 : G:0.808
_101 : G:1.000
_102 : T:0.999
_103 : K:0.779
_104 : L:0.669 V:0.330
_105 : T:0.915
_106 : V:0.999
_106a: L:0.968
_107 : G:0.728 R:0.205
_108 : Q:0.993
_109 : P:0.993

Table 6

Variable domains of the mouse lambda light chain

Determined consensus sequence

```
1 QAVVTQESA-LTTSPGETVTLTCRSSTGAVTTSN---YANWVQEKPDHLFTGLIGG
51 TNNRAPGVPARFSGSLIGDKAALTITGAQTEDEAIYFCALWYSNH-----WVFEGG
101 GTKLTVLGQP
```

(SEQ ID NO:6 shows the consensus sequence without introns)

Position Type and Frequency

1:	Q:0.966
2:	A:0.999
3:	V:1.000
4:	V:0.999
5:	T:1.000
6:	Q:1.000
7:	E:0.894 Q:0.105
8:	S:1.000
9:	A:0.999
10:	-:1.000
11:	L:0.999
12:	T:0.999
13:	T:0.999
14:	S:1.000
15:	P:0.999
16:	G:0.994
17:	E:0.655 G:0.344
18:	T:0.999
19:	V:0.999
20:	T:0.646 I:0.353
21:	L:1.000
22:	T:0.990

- 52 -

23 : C:1.000
24 : R:0.999
25 : S:0.999
26 : S:0.800 T:0.175
27 : T:0.888 S:0.112
28 : G:0.999
29 : A:0.999
30 : V:0.991
31 : T:1.000
31a: T:0.989
31b: S:0.925
31c: N:0.999
31d: -:1.000
31e: -:1.000
31f: -:1.000
32 : Y:0.999
33 : A:0.999
34 : N:0.990
35 : W:1.000
36 : V:0.919
37 : Q:1.000
38 : E:0.865 Q:0.135
39 : K:0.999
40 : P:0.990
41 : D:0.999
42 : H:0.999
43 : L:0.999
44 : F:0.999
45 : T:0.990
46 : G:0.999
47 : L:0.990
48 : I:0.982
49 : G:0.999
50 : G:0.984
51 : T:0.992
52 : N:0.609 S:0.343

- 53 -

53 : N:0.844
54 : R:0.999
55 : A:0.859 T:0.105
56 : P:0.999
57 : G:1.000
58 : V:0.999
59 : P:1.000
60 : A:0.632 V:0.367
61 : R:1.000
62 : F:1.000
63 : S:1.000
64 : G:1.000
65 : S:1.000
66 : L:0.999
67 : I:0.999
68 : G:1.000
69 : D:0.888 N:0.110
70 : K:0.999
71 : A:0.999
72 : A:0.999
73 : L:1.000
74 : T:0.999
75 : I:1.000
76 : T:0.999
77 : G:0.999
78 : A:0.928
79 : Q:1.000
80 : T:0.999
81 : E:1.000
82 : D:1.000
83 : E:0.657 D:0.343
84 : A:1.000
85 : I:0.615 M:0.385
86 : Y:1.000
87 : F:0.999
88 : C:1.000

- 54 -

89 : A:0.992
90 : L:0.999
91 : W:0.999
92 : Y:0.897
93 : S:0.895
94 : N:0.763 T:0.236
95 : H:0.929
95a: -:0.976
95b: -:0.999
95c: -:0.999
95d: -:0.999
95e: -:1.000
95f: -:1.000
96 : W:0.510 F:0.327 Y:0.107
97 : V:0.767 I:0.176
98 : F:1.000
99 : G:0.936
100 : G:0.841 S:0.159
101 : G:1.000
102 : T:0.996
103 : K:1.000
104 : L:0.549 V:0.451
105 : T:1.000
106 : V:1.000
106a: L:1.000
107 : G:1.000
108 : Q:0.874 X:0.126
109 : P:1.000

SEQUENCING PROTOCOL

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: BOEHRINGER MANNHEIM GMBH
- (B) STREET: Sandhofer Str. 116
- (C) CITY: Mannheim
- (D) COUNTRY: Germany
- (E) ZIP CODE: D-68305
- (G) TELEPHONE: 08856/60-3446
- (H) FAX: 08856/60-3451

(ii) TITLE OF INVENTION: Process for Modifying the Stability of Antibodies

(iii) NUMBER OF SEQUENCES: 10

(iv) COMPUTER-READABLE VERSION:

- (A) RECORDING MEDIUM: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30B (EPA)

(vi) DATA FOR SOURCE APPLICATION:

- (A) APPLICATION NUMBER: DE P 44 25 115.7
- (B) FILING DATE: 15-JUL-1994

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 113 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Glu	Val	Gln	Leu	Val	Glu	Ser	Gly	Gly	Gly	Leu	Val	Lys	Pro	Gly	Gly	
1				5					10					15		
Ser	Leu	Arg	Leu	Ser	Cys	Ala	Ala	Ser	Gly	Phe	Thr	Phe	Ser	Ser	Tyr	
			20					25					30			
Ala	Met	Ser	Trp	Val	Arg	Gln	Ala	Pro	Gly	Lys	Gly	Leu	Glu	Trp	Val	
		35					40					45				

Gly Trp Ile Tyr Asn Gly Gly Asp Thr Tyr Tyr Ala Asp Ser Val Lys
 50 55 60
 Gly Arg Phe Thr Ile Ser Arg Asp Thr Ser Lys Asn Thr Leu Tyr Leu
 65 70 75 80
 Gln Met Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ala Arg Gly Gly
 85 90 95
 Gly Gly Gly Tyr Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser
 100 105 110
 Ser

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 113 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION: 99...100
- (D) OTHER DATA: /product= "OTHER"
/note= "Xxa in locations 99 and 100 indicates - no amino acid -"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Glu Val Gln Leu Gln Gln Ser Gly Gly Glu Leu Val Lys Pro Gly Ala
 1 5 10 15
 Ser Val Lys Leu Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
 20 25 30
 Tyr Met His Trp Val Lys Gln Arg Pro Gly Lys Gly Leu Glu Trp Ile
 35 40 45
 Gly Arg Ile Asn Gly Ser Gly Gly Thr Asn Tyr Asn Glu Lys Phe Lys
 50 55 60
 Gly Lys Ala Thr Leu Thr Arg Asp Lys Ser Ser Ser Thr Ala Tyr Leu
 65 70 75 80
 Gln Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys Ala Arg Gly Gly
 85 90 95

Tyr Tyr Xaa Xaa Asp Tyr Trp Gly Gln Gly Thr Thr Val Thr Val Ser
 100 105 110

Ser

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 109 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
 1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Ser Ile Ser Ser Tyr
 20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
 35 40 45

Tyr Asp Ala Ser Asn Leu Glu Ser Gly Val Pro Ser Arg Phe Ser Gly
 50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
 65 70 75 80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Tyr Tyr Ser Leu Pro Tyr
 85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr
 100 105

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 108 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

Asp Ile Val Met Thr Gln Ser Pro Ala Ser Leu Ser Ala Ser Leu Gly
1           5           10           15
Glu Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Ser Val Ser Ser Tyr
20           25           30
Leu His Trp Tyr Gln Gln Lys Pro Gly Gln Ser Pro Lys Leu Leu Ile
35           40           45
Tyr Arg Ala Ser Asn Leu Ala Ser Gly Val Pro Asp Arg Phe Ser Gly
50           55           60
Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Val Glu Ala
65           70           75           80
Glu Asp Leu Ala Thr Tyr Tyr Cys Gln Gln Ser Asn Ser Tyr Pro Tyr
85           90           95
Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg
100          105

```

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 109 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:10
- (D) OTHER DATA: /product= "OTHER"
- /note= "Xxa in location 10 indicates - no amino acid -"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```

Gln Ser Glu Leu Thr Gln Pro Pro Ser Xaa Val Ser Val Ser Pro Gly
1           5           10           15
Gln Thr Val Thr Ile Ser Cys Ser Gly Asp Ser Leu Gly Ile Gly Tyr
20           25           30

```

```

Val Ser Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Lys Leu Val Ile
    35                                40                                45

Tyr Asp Asp Asn Lys Arg Pro Ser Gly Ile Pro Asp Arg Phe Ser Gly
    50                                55                                60

Ser Lys Ser Gly Asn Thr Ala Ser Leu Thr Ile Ser Gly Leu Gln Ala
    65                                70                                75                                80

Glu Asp Glu Ala Asp Tyr Tyr Cys Gln Ser Trp Asp Ser Ser Ser Val
    85                                90                                95

Val Phe Gly Gly Gly Thr Lys Leu Thr Val Gly Gln Pro
    100                                105

```

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 109 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION:10
- (D) OTHER DATA: /product= "OTHER"
- /note= "Xxa in location 10 indicates - no amino acid -"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```

Gln Ala Val Val Thr Gln Glu Ser Ala Xaa Leu Thr Thr Ser Pro Gly
1                                5                                10                                15

Glu Thr Val Thr Leu Thr Cys Arg Ser Ser Thr Gly Ala Val Thr Tyr
    20                                25                                30

Ala Asn Trp Val Gln Glu Lys Pro Asp His Leu Phe Thr Gly Leu Ile
    35                                40                                45

Gly Gly Thr Asn Asn Arg Ala Pro Gly Val Pro Ala Arg Phe Ser Gly
    50                                55                                60

Ser Leu Ile Gly Asp Lys Ala Ala Leu Thr Ile Thr Gly Ala Gln Thr
    65                                70                                75                                80

Glu Asp Glu Ala Ile Tyr Phe Cys Ala Leu Trp Tyr Ser Asn His Trp
    85                                90                                95

Val Phe Gly Gly Gly Thr Lys Leu Thr Val Gly Gln Pro
    100                                105

```

(2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single-stranded
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GGTAACACGT TCACCCAGTG ATACAGACAG AGAG

(2) INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single-stranded
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

CTGATACCAC GCCAGGTAGT TTTTCTGGTT ACC

(2) INFORMATION FOR SEQ ID NO: 9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single-stranded
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

ACCGCTACCG CTACCCGAGA AACGGTCCGG AACA

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single-stranded
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

CGGGTAAGAG TGGTCCTGTT GACAGTAGTA AAC

Claims

1. Process for preparing a stability-improved functional antibody, functional derivative, or fragment thereof in a eukaryotic or prokaryotic organism through transformation of the organism with an expression vector containing a recombinant gene, which codes for said antibody, derivative, or fragment, characterized in that

- a) the gene of at least one of the variable domains of the antibody, derivative, or fragment is compared with Consensus Tables 1-6 and the table is thereby selected which has the highest homology for this domain,
- b) at least one codon of an amino acid is replaced in the gene of this variable domain, namely
 - ba) if this amino acid is not listed in its position in said selected table, by a codon for one of the listed amino acids, and/or
 - bb) if this amino acid is listed in its position in the selected table, by a codon for one of the listed amino acids with a higher frequency,
- c) and the prokaryotic or eukaryotic organism with the thus modified gene is transformed and the antibody, the fragment, or derivative with the desired activity is expressed.

2. Process according to Claim 1, characterized in that a modified gene is used in which at least one codon for an amino acid is replaced, namely,

- a) in the gene of the variable domain of the human heavy chain,
 - aa) if this amino acid is not listed in its position in Table 1, by a codon for one of the listed amino acids, and/or
 - ab) if this amino acid is listed in its position in Table 1, by a codon for one of the listed amino acids with a higher frequency,

- b) in the gene of the variable domain of the mouse heavy chain
 - ba) if this amino acid is not listed in its position in Table 2, by a codon for one of the listed amino acids, and/or
 - bb) if this amino acid is listed in its position in Table 2, by a codon for one of the listed amino acids with a higher frequency,
- c) in the gene of the variable domain of the human kappa light chain
 - ca) if this amino acid is not listed in its position in Table 3, by a codon for one of the listed amino acids, and/or
 - cb) if this amino acid is listed in its position in Table 3, by a codon for one of the listed amino acids with a higher frequency,
- d) in the gene of the variable domain of the mouse kappa light chain
 - da) if this amino acid is not listed in its position in Table 4, by a codon for one of the listed amino acids, and/or
 - db) if this amino acid is listed in its position in Table 4, by a codon for one of the listed amino acids with a higher frequency,
- e) in the gene of the variable domain of the human λ light chain
 - ea) if this amino acid is not listed in its position in Table 5, by a codon for one of the listed amino acids, and/or
 - eb) if this amino acid is listed in its position in Table 5, by a codon for one of the listed amino acids with a higher frequency,
- f) in the gene of the variable domain of the mouse λ light chain
 - fa) if this amino acid is not listed in its position in Table 6, by a codon for one of the listed amino acids, and/or
 - fb) if this amino acid is listed in its position in Table 6, by a codon for one of the listed amino acids with a higher frequency.

3. Process according to Claim 1 or 2, characterized in that the antibody, fragment, or derivative is isolated from the organism and purified.
4. Process according to Claims 1 to 3, characterized in that the expression occurs in a prokaryotic or eukaryotic organism and the antibody, fragment, or derivative is formed functional in the cytosol.
5. Process according to Claims 1 to 4, characterized in that at least two codons are exchanged.
6. Process according to Claims 1 to 5, characterized in that in addition at least one of the disulfide-bond-forming cysteines in the antibody, fragment, or derivative is replaced by another amino acid.
7. Process according to Claim 6, characterized in that all disulfide-bond-forming cysteines are replaced.
8. Modification of the process according to Claims 1 to 5, characterized in that for the selective destabilization of an antibody, functional fragment, or derivative a more frequent amino acid is replaced by a less frequent amino acid.
9. Modified, stability-improved functional antibody, functional derivative, or fragment, obtained by the process according to Claims 1 to 7.
10. Stability-improved hybrid protein from a V_H and V_L domain of an antibody, obtained by the process according to Claims 1 to 7.
11. Recombinant disulfide-bond-free, functional, humanized, chimeric, nonhuman, or human antibody, derivative, or fragment assignable to β -lymphocyte expression.

12. Hybrid protein from a V_H and V_L domain of an antibody, which is linked by a intermolecular cystine bond, and in which the cysteines at positions 22 and/or 99 (heavy chain) 23 and/or 88 (light chain) are replaced by other amino acids and at least one other amino acid, preferably outside the CDR range, is replaced according to the process of Claims 1 to 5.
13. Use of an antibody obtained according to Claims 1 to 8 for the preparation of a medication, for example, for the treatment of cancer, autoimmune diseases, and infectious diseases.
14. Use of an antibody obtained according to Claims 1 to 7 in a diagnostic technique, for example, for the immunological analysis of substances binding to antibodies.
15. Process according to Claim 3, [characterized in] that the antibody, fragment, or derivative is folded under reducing conditions in vitro into the functional protein.
16. Use of an antibody, obtained according to Claims 1 to 7, as a catalytic antibody.

Fig. 1

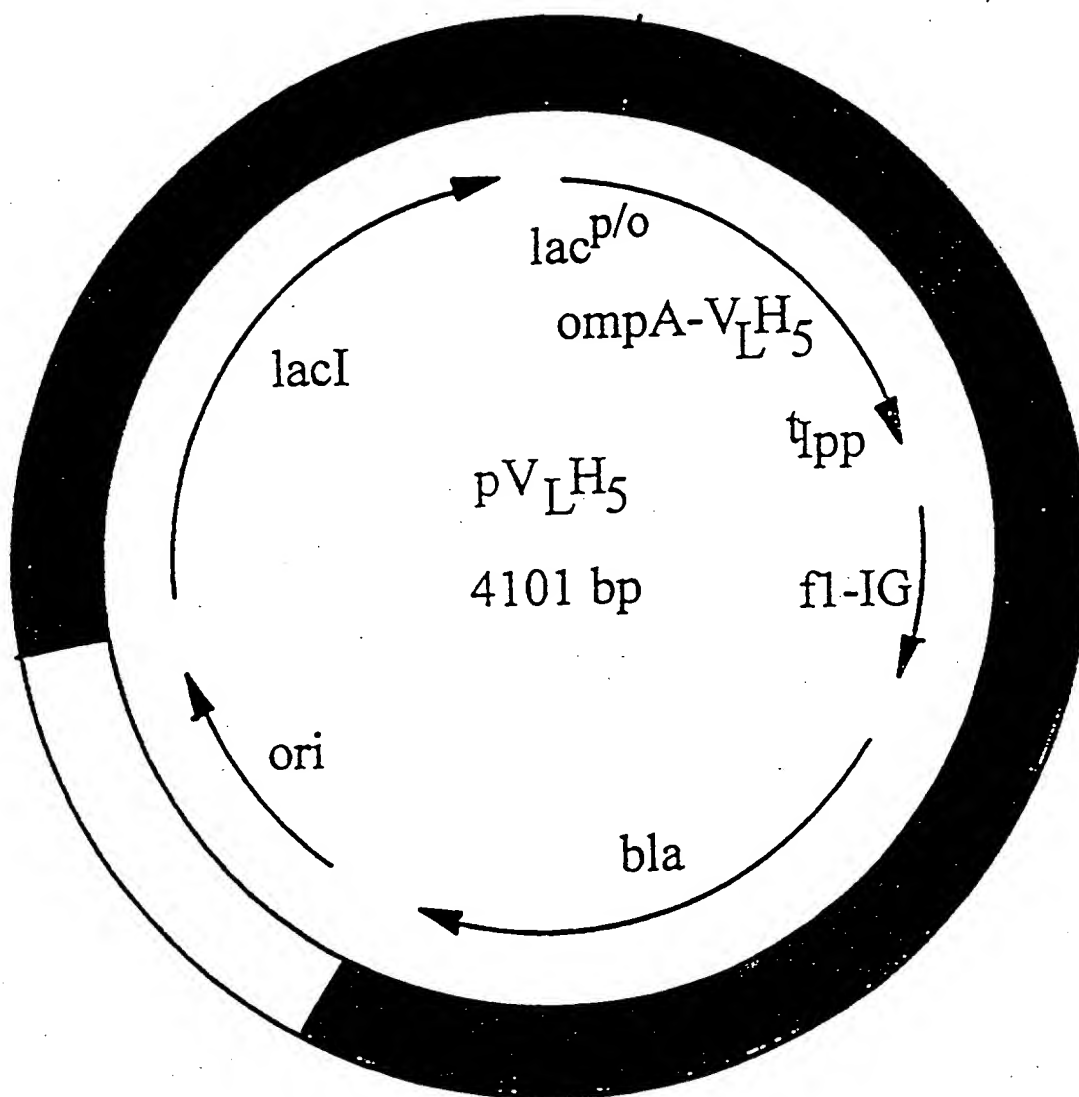


Fig. 2

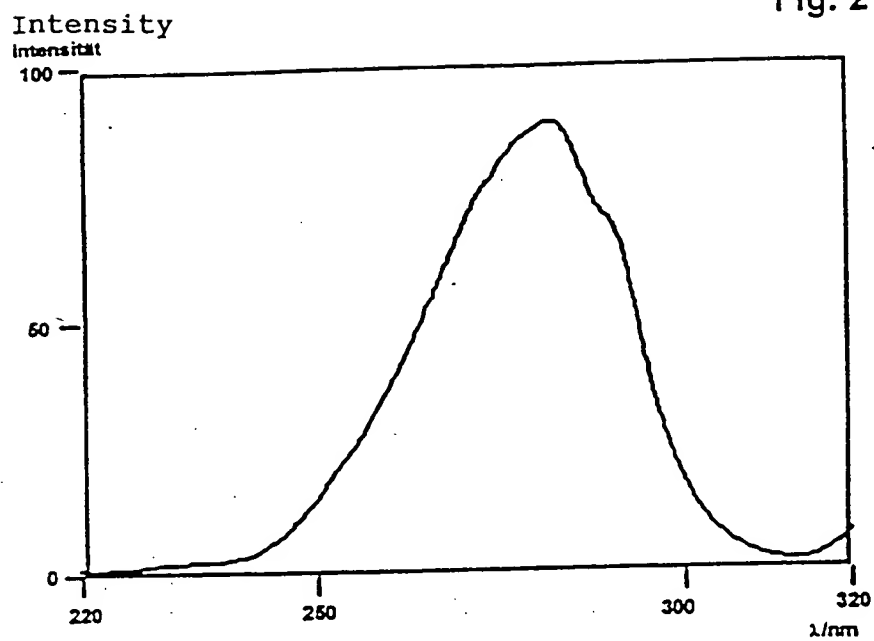


Fig. 3

